

Feature Importance を用いた株式投資のためのポートフォリオ評価の  
新たなリターン指標導出手法  
A Novel Dynamic Return Indicators of Portfolio Evaluation for Equity Investment Using Feature  
Importance

穴井 克和† Yoshikazu Anai † 岡田 龍太郎† Ryotaro Okada † 峰松 彩子† Ayako Minematu † 中西 崇文† Takafumi Nakanishi †

## 1. はじめに

近年, COVID-19 パンデミックをはじめ, それによる原油価格急落, ロシアによるウクライナ侵攻, インフレ率の急上昇, 仮想通貨市場の暴落, アメリカ債務上限問題, シリコンバレー銀行の破綻など, 世界にまたがる株式投資に関連する大きな出来事が頻発しており, 投資上級者にとっても見通しが読めない局面が多くなってきている. 特に, GameStop ショートスクイズは, アメリカ株式市場で Reddit の r/WallStreetBets コミュニティサイトの話題によって大量買いが発生し, ヘッジファンドが大きな損失を被り, 市場全体に混乱を招く事態が発生した. このように自然発生的に, もしくは人為的に市場が大きく混乱を招くことが多くなってきており, 急激な時代の変化によって, そのようなブラックスワンが増大すると予測されている. これらの事象は, 株式投資上級者であってもこれまでのテクニカル分析や財務指標分析のみでは限界があり, 狼狽売りが多く発生している. また, それらを防ぐために, 多様な種類のリスクに対して, 何を重視して判断するかを見極めることが重要となってきている.

一方で, 近年の AI, 機械学習の発展により, 株価予測に基づいた商品も販売されている. また, 説明可能な AI(Explainable AI, XAI)技術[1]により, 予測, 推定結果の理由を導出する手法も提案されつつある. 特に XAI 技術の中で Feature Importance[2]は, データセットから, 目的変数により貢献する説明変数を導出することを可能としている. Feature Importance 技術[2]を用いることでブラックスワンの事象発生時に, どのデータ, 指標に着目すべきかを選択し, その選択した特徴を用いてリターンの確信度を指標として導出し, 自身のポートフォリオの評価が可能になれば, ブラックスワンの事象発生した際に, リスクを最小にして, 分析が可能となり, 株式投資上級者にとってポートフォリオ組成の一助となりうると考えられる.

本稿では, Feature Importance を用いた株式投資のためのポートフォリオ評価の新たなリターン指標導出手法について示す. 本方式は, 株価変動を目的変数に置き, 財務データを説明変数とし, XAI 技術の一つである Feature Importance により, 株価変動を判断するのに有効な特徴を動的に抽出し, それぞれの企業に対する新たなリターン指標, 自身のポートフォリオに対する新たなリターン指標の導出手法を提案する. 本稿では, 目的変数を 3 年間の株価変動を対象としたが, それを 30 年の株価変動もしくは 3 秒の株価変動など違った時間窓で Feature Importance を求めることにより, 短期, 中期, 長期において重要な変数を動的に導出し, それぞれの企業, および自身のポートフォリオ

のリターン指標を計算することが可能であると考えられる. これにより, 上記で言及した, 様々な事象における株価変動の説明となる特徴を抽出し, リターン指標を計算することが可能となると考えられる. 本方式は, これまでの株式投資の評価, および自身のポートフォリオの評価において, 定点観測的な指標を用いていたのに対し, それぞれのリスクに応じた特徴を株価の変動から動的に選択をし, 評価が可能になると考えられる.

本稿の構成は以下の通りである. 2 節で関連研究について述べる. 3 節で本稿で提案する手法について述べる, 4 節で本稿で提案する手法の実現性を検証する実験を行い, 5 節でまとめる.

## 2. 関連研究

これまで, 企業の株価を評価する指標に関する研究は古くから多様に行われてきた. Dai[3]は, 上海証券取引所市場における企業の財務指標と株価の関係を回帰モデルを用いて調査し, 財務指標が株価に大きな影響を与えることが示唆され, 株価予測の理論的研究と実務の両面で役立つことが示唆している. Chang ら[4]は, ニューラルネットワーク, サポートベクターマシン(SVM), 混合データサンプリング(MIDAS)といった様々なデータマイニング技術を株価予測に用いており, その結果, 中期予測では MIDAS が, 短期予測では SVM がより精度が高いことを示している. Balasubramaniam[5]は, インド準備銀行指数, エコノミック・タイムズ指数, フィナンシャル・エクスプレス指数などの株価指数を評価し, これらの株価指数は, 株価の動きに伴う高いリスクとリターンを反映し, 証券価格と株式市場の動きを要約する指標として機能していると結論づけた. Chen ら[6]は, 対象企業の関連企業の情報をその株価予測に組み込むことを提案し, 実際の市場からの投資事実に基づいて, 関連するすべての企業を含むグラフを構築し, グラフに適用されたノード埋め込み法によって各企業の分散表現をグラフ畳み込みニューラルネットワークを介したジョイントモデルに基づいて, 株式予測を行っている. Bisarya[7]らは, LSTM を用いて株価予測を実現している. Gelb と Zarowin[8]は, 自主的な企業情報開示と株価の有益性の関係を検証しており, より多くの情報開示が将来の収益についてより情報に富んだ株価と関連しているという仮説を立て, それを発見している.

また, これまで株式のポートフォリオの指標に関する研究は古くから多様におこなわれてきた. Latané ら[9]は, 様々な市場インデックスを計算する際に使用されるウェイト

† 武蔵野大学データサイエンス学部 Department of Data Science, Musashino University

トシステムと再配分の手順について論じ、ポートフォリオ運用においてこれらの手順を理解することの重要性を強調し、様々な再配分方法の経時的なパフォーマンスを評価している。Maulana[10]は、株式ポートフォリオ分析について、単一インデックスモデルを適用して効率的なポートフォリオラインの $\beta$ 値を見つけ、投資家が最適なポートフォリオを形成するために必要な株式と資金の割合を決定できるようにすることを目的としており、検証をおこなっている。

Chou ら[11]は、ポートフォリオのパフォーマンスを評価する際には、リターンとリスクが重要な条件であり、シャープレシオは、ポートフォリオのリターンとリスクを同時に考慮するよく知られた評価戦略であるが、シャープレシオは平均線を使用してポートフォリオのリスクを評価するため、安定した上昇傾向にあるポートフォリオを高リスクと評価しやすいためリスクの低い安定した上昇傾向のポートフォリオを評価し、単位日次リスクあたりの期待リターンが高いポートフォリオを特定することが可能なトレンド比を使用してこの問題に対処している。さらに、限られた時間内にポートフォリオを効果的に最適化するために、Global-best guided Quantum-inspired Tabu Search algorithm with Not-gate (GNQTS) を使用し、13 種類のスライディングウィンドウを使用してデータをトレーニングおよびテストし、トレンド比の高い適切な期間を特定している。Chang ら[12]は、Process Capability Indices (PCI) を適用して、新しいパフォーマンス評価方法を提案し、また investment satisfied capability index (ISCI) 指数を利用して個々の株式のパフォーマンスを迅速に評価し、投資満足度の達成につながる株式を選択できる手法を提案している。また、移動間隔ウィンドウを使用した particle swarm optimization (PSO) アルゴリズムを適用して、このポートフォリオ内の株式の最適な投資配分を見つけることを可能としている。Vasiani ら[13]は、priority index method と遺伝的アルゴリズムの両方を適用して、収益の観点から株式ポートフォリオを最適化する手法を提案している。またポートフォリオの最適化は文献[14]が詳しい。また、Azizah ら[15]は、株価の評価に使用される single index model と、投資する資産の割合を決定するためにラグランジュ乗数法を使用して開発された定式化最適化モデルについて示している。

これらの手法は特定の特徴量を用いた、その企業の株価の評価、およびポートフォリオの評価を実現している。それに対し、本研究では、Feature Importance を導出することにより、株価変動を判断するのに有効な特徴を動的に抽出し、それぞれの企業に対する新たなリターン指標、自身のポートフォリオに対する新たなリターン指標を実現している。またポートフォリオの最適化については、上記で紹介した手法と組み合わせて使うことができると考えられる。正確な予測を提供するという点で、人工知能および周波数領域モデルよりも優れたパフォーマンスを示した。この研究では、株式リターンの予測に機械学習モデルを有用な手段として適応できることの可能性を示唆している。

### 3. 全体像

図 1 に本方式の全体像を示す。本方式では大きく分けて、企業データベースと、財務指標特徴量抽出機能、アクティブ・ポートフォリオ組成機能、リターン指標作成機能、ポートフォリオ全体評価機能の 4 つの機能で構成される。

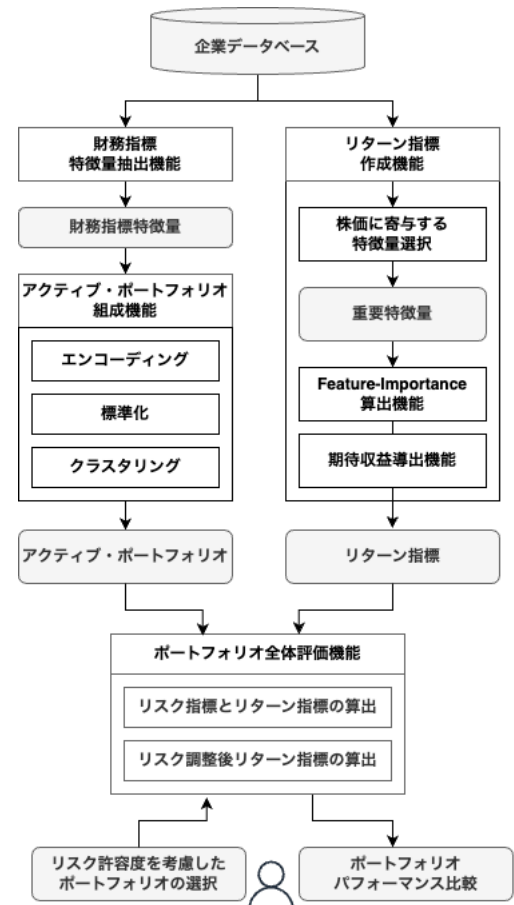


図 1 提案方式の全体像

本研究では、XAI 技術の一つである Feature-Importance を用いて期待収益を導出する。導出した期待収益の他に、複数の評価指標を用いることで、ポートフォリオ全体のパフォーマンス比較例を示す。また、本研究で導出する期待収益の有効性を検証するため、クラスタリングを用いてポートフォリオを組成する。ポートフォリオの組成には様々な手法が存在し、クラスタリング以外の手法も考えられたが、組成のしやすさの観点からクラスタリングを採用した。具体的には、財務指標と株価データをクラスタリングして、形成されたクラスターをポートフォリオに用いる。これにより、財務指標や株価データという実績に基づいたグループが形成されることから、業種別や業界別に近いアクティブ・ポートフォリオとして期待される。以上により、組成されたポートフォリオを本研究に用いて、パフォーマンス比較例を示す。

### 3.1 企業データベースの構築

企業データベースの構築にあたり、各社の財務指標および株価データを取得した。データの収集には、金融庁の電子開示システム「EDINET」[16]、企業分析サイト「パフェットコード」[17]を使用した。対象企業は、2024 年 4 月 9 日時点の東京証券取引所に上場している時価総額ランキング上位 400 社とした。収集したデータをデータベースに格納し、企業データベースを構築した。

表 1 抽出した特徴量

指標名	性質
売上 CAGR5 年平均	成長性
営業利益 CAGR5 年平均	成長性
EPS CAGR5 年平均	成長性
投資 CF	成長性
営業利益率	資本効率
ROE	資本効率
棚卸資産回転日数	資本効率
PBR	バリュエーション
自己資本比率	財務安定性
流動比率	財務安定性
有利子負債 / 総資産	財務安定性
営業 CF	財務安定性
財務 CF	財務安定性
3 年株価上昇率	株価

表 2 エンコーディング後の CF

会社名	営業 CF	投資 CF	財務 CF
ローゼ	0	0	1
SCREENHD	1	0	0
アドバンテスト	1	0	0
川崎汽船	1	0	0
IHI	1	0	0

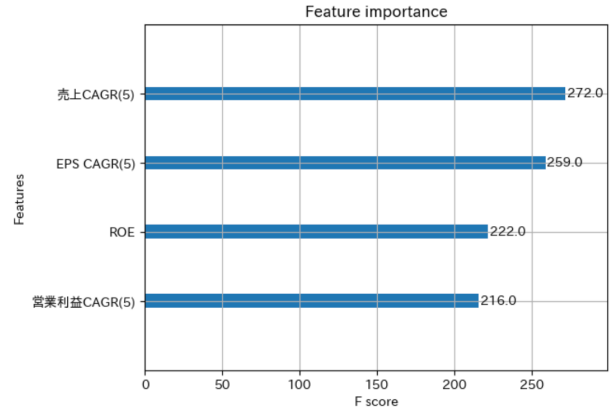


図 2 Feature-Importance

### 3.2 財務指標特徴量抽出機能

本節では、3.1 で構築した企業データベースから分析に使用する特徴量を抽出する。特徴量は無数に考えられるが、その中で重要と思われる特徴量を選んだ。その際には、一般的な財務知識における重要性を考慮したほか、特徴量間の相関を計量し、相関が互いに高い特徴量は片方のみを採用するなどして特徴量を厳選した。抽出した特徴量は、3.3.3 節で実施するクラスタリングに使用する。さらに、特徴量をユーザが見て判断しやすいように、特徴量の性質を成長性、資本効率、バリュエーション、財務安定性、株価の 5 つのカテゴリに分類して付与した。抽出した特徴量とその性質を表 1 に示す。

### 3.3 アクティブ・ポートフォリオの組成

#### 3.3.1 エンコーディング

抽出した特徴量には、営業 CF（キャッシュフロー）、投資 CF、財務 CF が含まれる。これらの指標は、会社の規模によって値の大小が変わることから、同じスケールでの会社同士の比較には向かない。以上のことから、CF の値がプラスの時は「1」、値がマイナスの時は「0」のデータへと変換した。変換したデータの一例を表 2 に示す。

#### 3.3.2 標準化

各特徴量はデータの上限值や下限値、平均値などが異なるため、その影響を揃えるために、すべての特徴量に対して標準化を行う。

#### 3.3.3 クラスタリング

前節までの前処理を施した財務指標を特徴量に用いて、k-means によりクラスタリングした。クラスタ数は、企業数や含まれている企業群のバリエーションを考慮して実験の際に決定する。

形成された各クラスタは財務指標と株価が似ている会社で構成される。そのため、業種別アクティブ・ポートフォリオを用いた投資商品に近いパフォーマンスが期待されると考える。ここでは、各クラスタをポートフォリオとして用いることとする。

### 3.4 リターン指標作成機能

本節では、ポートフォリオの期待収益を導出するための、リターン指標作成機能について述べる。期待収益とは、ポートフォリオが生み出すと予測されるリターンを表す。リターン指標作成機能は、株価に寄与する特徴量選択機能、Feature-Importance 算出機能、期待収益導出機能の 3 つの機能からなる。

株価に寄与する特徴量の選択機能では、精度向上を目的として、特徴量の絞り込みを行う。Feature-Importance 算出機能では、機械学習の分類アルゴリズムである XG-Boost における Feature-Importance を算出する。期待収益導出機能では、Feature-Importance の値を用いた期待収益を求める式を定義し、その式を用いて期待収益を導出する。

#### 3.4.1 株価に寄与する特徴量の選択機能

本節では、3 年株価上昇率を目的変数として、分類に用いる特徴量の選択を行う。精度向上を目的として、表 1 に定義した 3 年株価上昇率を除いた 13 個の特徴量の中から、株価への影響が大きい特徴量を選択し絞り込みを行う。ここでは機械学習の分類アルゴリズムである XG-Boost を使用し、さらに機械学習モデルに対する特徴量の絞り込み手法である scikit-learn の selectKBest を用いることで特徴量の絞り込みを行う。教師データとして 3.1 節で構築した企業データベースで対象とした 400 社のデータを用いた。機械学習モデルに対する入力としては、表 1 の特徴量のうち、3 年株価上昇率を目的変数とし、残りの 13 個の特徴量を説明変数とした。データをラベル付けするために、3 年株価上昇率の平均値より大きい値を「1」、平均値より低い値を「0」とした。

分類モデルの構築後に、selectKBest を用いて特徴量の絞り込みを行った。特徴量の選択数を  $m$  とする。 $m=4$  としたとき、選択された特徴量は以下の通りとなった。

1. 売上 CAGR5 年平均
2. EPS CAGR5 年平均
3. ROE
4. 営業利益 CAGR5 年平均

### 3.4.2 Feature – Importance算出機能

前節では、13 個の特徴量を説明変数として、3 年株価上昇率を目的変数として XG-Boost を用いてモデルを構築し、その中から、予備実験により精度が高かった  $m$  個 (ここでは  $m=4$ ) の特徴量を重要なものとして選択した。本節では、選択された  $m$  個の説明変数を用いて再度 XG-Boost を用いてモデルを構築し、その場合における特徴量の重要度を算出する。

XG-Boost で構築されたモデルには、Feature-Importance という予測結果に対する特徴量の重要度を表現した指標が存在しており、この Feature-Importance の値を正規化して特徴量の重みとして採用し以後の計算に用いる。図 2 に正規化前の Feature-Importance の値を示す。

正規化前の Feature-Importance を  $w'_i$  としたとき、正規化後の Feature-Importance である  $w_i$  を以下の式で表す。ここで  $m$  は絞り込み後の特徴量の数を表す。ここでは正規化として、すべての Feature-Importance の値の和で各 Feature-Importance の値を割っている

$$w_i = \frac{w'_i}{\sum_{i=1}^m w'_i}$$

### 3.4.3 期待収益導出機能

前節で求めた Feature – Importance  $w_i$  と、その特徴量に対応する正規化後の特徴量の実際の値  $x_i$  を 2 乗したものをそれぞれ掛け合わせて総計をとることにより期待収益  $e_j$  を導出する。ここで  $j$  は企業を識別するインデックスである。 $x_i$  を 2 乗するのは、 $x_i$  は、 $-1 \sim 1$  の範囲で正規化された値であるためマイナスの値をとることがあるが、ここでは符号は考慮せず値の大きさにのみ注目するからである。期待収益  $e_j$  を求める式を以下に示す。

$$e_j = \sum_{i=1}^m x_i^2 \cdot w_i$$

## 3.5 ポートフォリオ全体評価機能

ポートフォリオのパフォーマンスを評価し、比較する指標として、リスク指標  $\beta$ 、リターン指標  $\bar{R}$ 、リスク調整後リターン  $RR$  を導出する。

### 3.5.1 リスク指標とリターン指標の算出

ポートフォリオのパフォーマンスを評価する指標のうちリスクを評価する指標をリスク指標  $\beta$  として定義する。 $\beta$  の算出には  $\beta$  値を使用する。 $\beta$  値とは、個別証券または、ポートフォリオの収益が証券市場全体の動きに対してどの程度敏感に反応して変動するかを示す数値である。 $\beta$  値は、投資家の間で一般的に用いられる指標である。本節で使用する  $\beta$  値の算出には以下の式を用いた。

$$\beta = \frac{\text{Cov}(R_A, R_T)}{\text{Var}(R_T)}$$

$R_A$  は A 社の株価の月次リターン、 $R_T$  は TOPIX の月次リターンであり、 $\text{Cov}(R_A, R_T)$  は、 $R_A$  (A 社の月次リターン) と  $R_T$  (TOPIX の月次リターン) の共分散を表す。 $\text{Var}(R_T)$  は TOPIX の月次リターンの分散を表す。

表 3 各クラスタの特性の評価

クラスタ	特性
0	成長性に優れており、財務的に安定している
1	全体的に値は低く、有利子負債 / 総資産が高い。
2	成長性に優れており、バリュエーション (以下、PBR という) が低い。
3	財務の安定に特化している

本研究では、リスク指標として、会社ごとの  $\beta$  値とその会社の投資比率を乗じた値の全体の平均値を採用した。投資比率は実際には任意に設定可能であるが、本研究では全ての銘柄を均等に割った数字とした。個別銘柄の  $\beta$  値を  $\beta_j$ 、投資比率を  $r_j$  とする。 $n$  は企業の総数を表す。リスク指標  $\bar{\beta}$  を求める式を以下に定義する。

$$\bar{\beta} = \frac{1}{n} \sum_{j=1}^n (\beta_j \cdot r_j)$$

リターン指標  $\bar{R}$  は、3.4 節で作成した期待収益  $e_j$  と投資比率  $r_j$  を乗じた値のクラスタ全体の平均値とする。リターン指標  $\bar{R}$  を求める式を以下に定義する。

$$\bar{R} = \frac{1}{n} \sum_{j=1}^n (e_j \cdot r_j)$$

### 3.5.2 リスク調整後リターン指標の算出

リスクとリターンを考慮した総合的な指標として、リスク調整後リターン指標を算出する。ポートフォリオ全体のリスク指標を  $\beta$ 、ポートフォリオ全体のリターン指標  $\bar{R}$  としたとき、リスク調整後リターン指標を  $RR$  とし、以下の式で定義する。

$$RR = \frac{\bar{R}}{\beta}$$

## 4. 実験

本節では、提案手法の有効性を検証するために、具体的な企業のデータを用いてポートフォリオを作成し、さらにそれぞれのポートフォリオについて指標を提示することで、ユーザがポートフォリオを選択するための補助となっていることを検証する。

4.1 節では実験環境を示す。4.2 節では実験 1 として、クラスタリングを用いたポートフォリオ組成の検証を行う。

4.3 節では実験 2 として、提案した指標である期待収益の有効性の検証を行う。4.4 節では実験 3 として、評価指標を用いたポートフォリオ分析の検証を行う。4.5 節では実験全体の考察を行う。

### 4.1 実験環境

企業データベースの構築にあたり、各社の財務指標および株価データを取得した。データの収集には、金融庁の電子開示システム「EDINET」[16]、企業分析サイト「バフェットコード」[17]を使用した。対象企業は、2024 年 4 月 9 日時点の東京証券取引所に上場している時価総額ランキング上位 400 社とした。

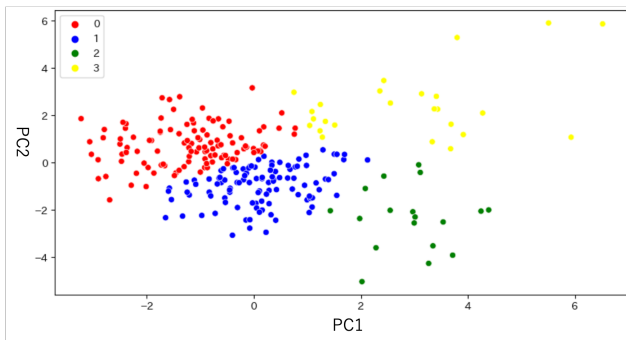


図 3 企業のクラスタリング結果の主成分分析による可視化

プログラムの構築には Python3.11.4 言語を用いた。3.4 節のリターン指標作成機能における機械学習アルゴリズムには scikit-learn1.3.2 を用いた。csv ファイルの操作には pandas2.0.3、データの可視化には matplotlib3.7.1 を用いた。

#### 4.2 実験 1: クラスタリングを用いたポートフォリオ組成の検証

##### 4.2.1 実験目的

企業群のクラスタリングを行った際に、それが適切なクラスタに分かれているのかを検証する。まず、4.2.2 節では、クラスタ数の決定を行う。4.2.3 節では、クラスタリング結果の検証を行い、各クラスタがどのような特徴を持っているのかについて確認する。4.2.4 節では、クラスタリングが適切であることを検証するため、クラスタにその特性を表す名前を命名し検証する。

##### 4.2.2 クラスタ数の決定

企業データベースに含まれる各企業について、表 1 に示す特徴量を標準化したベクトルを基準にクラスタリングを行う。クラスタリングには k-means を用いた。ここでは、 $k=4$  とした。全データのクラスタリング結果に対して主成分分析を行って 2 次元に写像した各企業のデータを図 3 に示す。図 3 を見る限り、3 つあるいは 4 つに分けるのが望ましいと考えたが、今回は企業の特性を細かく分けて考えることを目指し、クラスタ数を 4 とした。

##### 4.2.3 クラスタリング結果の検証

クラスタリングによって得られたポートフォリオ(クラスタ)について、その特徴量の傾向を可視化し、各ポートフォリオの特徴について考察する。各ポートフォリオの平均値を標準化したレーダーチャートを図 4 に示す。図 4 の結果を基に、4 つのポートフォリオの特性を定性的に記述したものを表 3 に示す。各ポートフォリオの特性は違いははっきりしており、クラスタリングは有効に行っていると判断できる。

##### 4.2.4 クラスタへの命名の検証

クラスタリングが適切であることを検証するため、ポートフォリオに特性を表す名前を命名し、検証する。以下に命名した結果を示す。

- クラスタ 0: 財務盤石グロース企業  
成長性に優れており、財務的に安定している。PBR も高く市場ではグロース銘柄として期待される。
- クラスタ 1: 低成長重長大企業

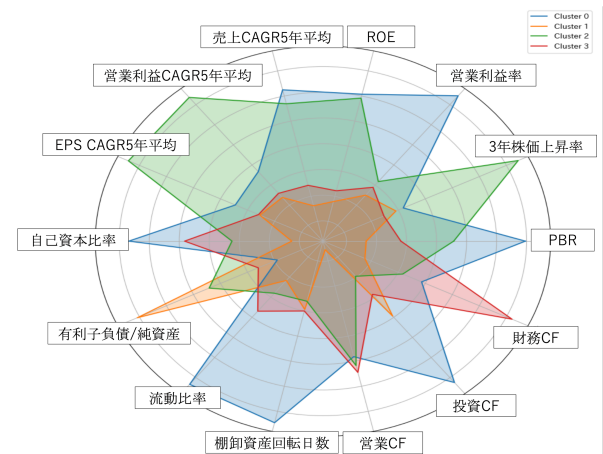


図 4 平均値を使用したレーダーチャート

指標の結果から財務的な安定感はない。しかし、クラスタリングされた企業を分析すると、不動産や自動車メーカー、化学メーカーといった業種が多いことがわかった。これらの業種は、工場や不動産を多く抱えるため、建設費や購入費用が高く、キャッシュで工面することが難しい。そのため、他の業種と比べると財務的な部分で悪く見えてしまう傾向がある。以上の理由から、負債の少ない業種と単純比較するのは適切ではないと判断し、財務が悪いと命名するよりも、クラスタリングされた企業の業種を反映した命名が適切と判断した。

##### 3. クラスタ 2: 優等生企業

成長性に優れており、PBR が低いことから市場ではバリュー銘柄と扱われる。

##### 4. クラスタ 3: 財務特化企業

成長性は欠けるが、財務の安定に特化している。

#### 4.3 実験 2: 期待収益の有効性の検証

##### 4.3.1 実験目的

本研究で提案した指標である期待収益が、既存の指標より優れた指標であるかどうかを検証する。具体的には、期待収益は将来の株式のリターンを表す指標であるため、株価の変動に対する影響度をリターン指標としての有効性の観点から評価する。

##### 4.3.2 実験結果

期待収益の有効性を検証するために、3 年の株価上昇率に対する影響を調べた。また、表 1 に示した特徴量のうち、3 年の株価上昇率との相関が最も強かった EPS CAGR5 年平均と、本研究で導出した期待収益の 2 つが、3 年の株価上昇率にどの程度影響を与えるのかを調べる。検証には ANOVA 分析を用いた。分析結果を表 4 に示す。

##### 4.3.3 考察

EPS CAGR5 年平均と期待収益の 2 つが、3 年の株価上昇率にどの程度影響を与えるのかを調べた。総じて、期待収益が 3 年の株価上昇率に影響する結果となった。財務データという株価を裏付ける指標から期待収益を算出できたことが、従来の金融理論と一線を画す点である。

表 4 ANOVA 分析の結果

	平方和	自由度	F 値	p 値
EPS CAGR 5 年平均 期待収益 率 残差	288.73	1	0.08	0.77
	102918	1	29.41	0.000001
	909805	260		

#### 4.4 実験 3 : 評価指標を用いたポートフォリオ分析の検証

##### 4.4.1 実験目的

各ポートフォリオはそれぞれが個性を持ったポートフォリオであるため、ユーザがポートフォリオを簡単に比較できるように、本研究で算出した評価指標を用いて分析する。評価指標には、3.5 節で算出したリスク指標 $\beta$ 、リターン指標 $\bar{R}$ 、リスク調整後リターン指標 $RR$ を用いた。

##### 4.4.2 実験結果

ポートフォリオの評価結果を正規化したものを表 5 に示す。リターン指標 $\bar{R}$ では「優等生企業」が最も高い結果を示した。リスク指標 $\beta$ では、「財務特化企業」が最も低い結果を示した。リスク調整後リターン指標 $RR$ では優等生企業が最も高い結果を示した。

##### 4.4.3 考察

リターン指標 $\bar{R}$ では、命名通り「優等生企業」が圧倒する結果となった。リスク指標 $\beta$ では、最も低い結果を示した「財務特化企業」は、財務の安定に特化しているため、他のポートフォリオよりリスクを取っていない特徴を持っている。そのため、市場の動きに対して価格が安定していることを示唆する。リスク調整後リターン指標 $RR$ では、「優等生企業」が最も高い結果となり、リターン指標 $\bar{R}$ の高さが影響する形となった。

#### 4.5 実験全体の考察

本節では、実験全体について考察する。考察を以下に述べる。

実験 1 では、ポートフォリオの組成のしやすさの観点から、クラスタリングによって形成されたクラスタをポートフォリオとして用いた。各企業を財務指標と株価データから効率的にグループ化することで、それぞれが個性あるポートフォリオが組成された。今後の発展としては、より運用を想定したポートフォリオを組成できるように、形成されたクラスタの中から、さらに詳細に財務指標を分析することで、より実践を想定したポートフォリオの組成の実現が考えられる。

実験 2 では、ポートフォリオのパフォーマンス比較例として、XAI 技術の一つである Feature-Importance を用いて期待収益を導出した。導出した期待収益と株価の変動に対する影響の観点から、既存の財務指標と比較した分析では、期待収益がより優れたパフォーマンスを示した。財務データという株価を裏付ける指標から期待収益を導出できたことが、従来の金融理論と一線を画す点である。また、リスクとリターンを複合的に分析することで、ポートフォリオの特性を把握する手助けとなる。

表 5 各ポートフォリオに対する評価指標の値

	財務盤石 グロース 企業	低成長重長 広大企業	優等生 企業	財務特化 企業
$\bar{R}$	0.54	0	1	0.037
$\beta$	1	0.063	0.98	0
$RR$	0.66	0	1	0.073

実験 3 では、本研究で算出したリターン指標 $\bar{R}$ 、リスク指標 $\beta$ 、リスク調整後リターン $RR$ を用いてポートフォリオのパフォーマンス比較例を示した。各指標はポートフォリオの特性を捉えており、特に 4.2.4 で実施した、クラスタに付与した名称を反映する結果を示した。これは、本研究で算出した評価指標が、ポートフォリオのパフォーマンスの比較およびポートフォリオの特性の理解の一助となることを示唆している。

#### 5. おわりに

本稿では、Feature Importance を用いた株式投資のためのポートフォリオ評価の新たなリターン指標導出手法を示した。XAI 技術である Feature-Importance を用いて、株式の期待収益を導出した。従来手法では、過去の株式の変動から期待収益を導出していたが、本研究では株価に寄与する財務指標を用いて期待収益を導出した点が、これまでの金融理論と一線を画す点である。また、株価に寄与する財務指標を XAI 技術を用いて定量的に算出できたことは、金融分野における XAI 技術の有効性を示す一例となった。

本研究が提案する Feature-Importance を用いた期待収益の導出手法が実務の場面で採用されることで、ポートフォリオがもたらすリターンを見積もることにつながり、結果として将来に対する投資戦略の一助となり得る。これにより、ポートフォリオを組成する運用担当者および個人投資家など、運用を行う全ての人々が株式リターンを見積もりに役立てることができると期待される。今後の展望として、目的変数に採用した 3 年の株価上昇率だけでなく、期間を変えた 5 年、10 年といった中期から長期における Feature-Importance を比較することで、投資家の求める期間に対応する期待収益を導出することが期待される。

#### 参考文献

- [1] Gunning David, Aha David, "DARPA's explainable artificial intelligence (XAI) program," in AI Magazine, Vol. 40, No. 2, pp. 44-58, (2019).
- [2] Tian Jilun, Jiang Yuchen, Zhang Jiushi, Wang Zhenhua, Rodriguez-Andina Juan J., Luo Hao, "High-performance fault classification based on feature importance ranking-XgBoost approach with feature selection of redundant sensor data," in Current Chinese Science, Vol. 2, No. 3, pp. 243-251, (2022).
- [3] Dai Li-ping, "An Empirical Study on Relationships between Financial Index and Stock Pricing," in Journal of Yunnan University of Finance and Economics, (2004).
- [4] Chang To-Han, Wang Nientsu, Chuang Wen-Bin, "Stock price prediction based on data mining combination model," in Journal of Global Information Management (JGIM), Vol. 30, No. 7, pp. 1-19, (2021).
- [5] Balasubramaniam C. S., "Indexes of Ordinary Share Prices-An Evaluation," in Artha Vijnana: Journal of The Gokhale Institute of Politics and Economics, pp. 552-564, (1980).
- [6] Chen Yingmei, Wei Zhongyu, Huang Xuanjing, "Incorporating corporation relationship via graph convolutional neural networks for stock price prediction," in Proceedings of the 27th ACM International

- Conference on Information and Knowledge Management, pp. 1655-1658, (2018).
- [7] Bisarya Udbhav, Parekh Vishwas, Bhattacharjee Shrutilipi, "Stock Price Prediction Using Corporation Network and LSTM," in 2022 2nd International Conference on Intelligent Technologies (CONIT), IEEE, pp. 1-6, (2022).
- [8] Gelb David S., Zarowin Paul, "Corporate disclosure policy and the informativeness of stock prices," in Review of Accounting Studies, Vol. 7, pp. 33-52, (2002).
- [9] Latané Henry A., Tuttle Donald L., Young William E., "Market indexes and their implications for portfolio management," in Financial Analysts Journal, Vol. 27, No. 5, pp. 75-85, (1971).
- [10] Maulana Yasir, "Stock investment portfolio analysis with single index model," in Indonesian Journal Of Business And Economics, Vol. 3, No. 2, (2020).
- [11] Chou Yao-Hsin, Lai Yun-Ting, Jiang Yu-Chi, Kuo Shu-Yu, "Using trend ratio and GNQTS to assess portfolio performance in the US stock market," in IEEE Access, Vol. 9, pp. 88348-88363, (2021).
- [12] Chang Jui-Fang, Shi Peng, "Using investment satisfaction capability index based particle swarm optimization to construct a stock portfolio," in Information Sciences, Vol. 181, No. 14, pp. 2989-2999, (2011).
- [13] Vasiani V. D., Handari Bevina D., Hertono G. F., "Stock portfolio optimization using priority index and genetic algorithm," in Journal of Physics: Conference Series, Vol. 1442, No. 1, pp. 012031, (2020).
- [14] Best Michael J., "Portfolio optimization," CRC Press, (2010).
- [15] Azizah E., Rusyaman E., Supian S., "Optimization of investment portfolio weight of stocks affected by market index," in IOP Conference Series: Materials Science and Engineering, Vol. 166, No. 1, pp. 012008, (2017).
- [16] "EDINET", <https://disclosure2.edinet-fsa.go.jp/WEEK0010.aspx> (最終閲覧日 2024-04-09)
- [17] "バフェットコード", <https://www.buffett-code.com/>, (最終閲覧日 2024-04-09)